STAT

HYPHENLESS JUSTIFICATION

typesetting breakthrough

by GEORGE E. KUNKEL and TILMON H. MARCUM

While an increasingly large number of computer and printing organizations continue to struggle with the hyphenation problem in computer produced text, a small team of specialists in the Central Intelligence Agency has developed a novel approach which is amazing in its simplicity and which may make the troublesome end-of-line hyphen a thing of the past in computerized photocomposition.

The success or failure of present computerized book composition systems hinges largely upon the computer's ability to provide correct end-of-line word divisions required for justification. Several years and hundreds of thousands of dollars have been spent in attempting to develop dictionaries and logic which will attain the necessary accuracy required for high quality composition. Existing computer systems cannot yet equal the accuracy of the average keyboard operator in hyphenating end-of-line words although a fairly high degree of accuracy is obtained in some cases. Accuracy is needed in computer composition because of the problem of correcting errors and the constant demand for greater speeds.

With the present state of the art, the printer must choose between a system obtaining high accuracy with proportionately higher computer costs or a less sophisticated system providing a lesser degree of accuracy with the attendant error cost factor. Any system permitting errors poses the problem of new errors being introduced when corrections are required.

The idea of eliminating the end-of-line hyphen is not new. Some newspapers and printers now justify text without hyphens by the use of excessive interword spacing and fixed letter spacing or a combination of both. Some publications appear with an unjustified single or multi-column format. Unfortunately neither of these systems provides acceptable typographic quality for book production.

A primary rule in typesetting is to avoid hyphens wherever possible since they destroy continuity in reading. It is obvious, then, that the real need is for a simple plan which eliminates the end-of-line hyphen without sacrificing typographic quality. Uncle Sam's team feels that they have solved this problem by the use of a variable set size technique on a line-for-line basis.* In simpler terms, this means that a sort of "coefficient of expansion" or contraction is applied to the proportional spacing between characters in each line of text. The line is thus expanded and contracted without destroying the proportional values of the individual character as is the case with fixed letter-spacing. The variation in appearance of the lines of text produced by this method is sufficiently subtle to remain unnoticed by the average reader. The ability to expand and contract provides sufficient latitude in justification so that the need for end-of-line word division is rare. The computer programming and processing which is required for variable set size justification is many times simpler and less costly than that required for end-of-line word division and hyphenation.

In a unit font each character is assigned a proportional unit value, which is valid regardless of the set size. The specified line measure can then be stated in units for each set size. Therefore the computer can accumulate units as it passes through four overlapping (set sizes) zones of justification in which an interword space may be selected for the end of line. This simple logic for line justification thus eliminates the requirements for stored dictionaries for word hyphenation, programmed logic approaches, or a combination of the two, prefix and suffix tables, and stripping and reconstituting routines.

The Central Intelligence Agency prints in its own facilities a number of high quality book-type publications. These publications are set in type on Intertype and Mono-type machines and are printed by offset from plates made from positive Mylar proofs pulled directly from the type. For the past five years the majority of this composition has been done on Model F4 Intertype machines operated by TTS tape.

About a year ago, the agency printing engineers and computer specialists began studying the possibility of doing page composition by photocomposing from computer-prepared tape. This study developed the feasibility of this type of composition and a thorough investigation was made of all available photocomposing equipment. After a great deal of study a Model 513 Photon has been obtained for this composition. A primary reason for this selection was the 513's capability of changing set sizes from codes inserted into the computer-prepared tape.

The printing engineers and computer specialists began the job of programming an IBM 1410 computer for this typesetting job. Agency personnel had talked with other

*Set size determines the overall horizontal dimension of the space assigned to a character—it includes space for separation from adjacent letters. Set size also expresses the relationship of one type font to another. The size relationship of one letter to another is expressed in "units of relative value," based on the size of the em (18/18). Therefore the horizontal dimension for a letter is determined:

$$URV \times SS = \text{Horizontal dimension}$$

$$\text{Point Size} \left(\frac{1"}{72} \right) = \text{Vertical dimension}$$